
PERBANDINGAN METODE NAIVE BAYES CLASSIFIER DENGAN METODE DECISION TREE (C4.5) UNTUK MENGANALISA KELANCARAN PEMBIAYAAN (Study Kasus : KSPPS / BMT AL-FADHILA)

Triowali Rosandy

Informatics and Business Institute Darmajaya
Email: Triowali.andy@gmail.com

ABSTRAK

Salah satu algoritma klasifikasi yang sering digunakan dan mendapat banyak perhatian para peneliti dalam memprediksi pembiayaan bermasalah pada perbankan adalah *Naive Bayes Classifier* dan *Decision Tree (C4.5)*. Tetapi kelemahan yang dihadapi pada kedua algoritma tersebut adalah lamanya waktu dan tingkat akurasi prediksi yang digunakan untuk melakukan prediksi. Masalah ini juga menjadi perhatian banyak peneliti untuk memperbaikinya agar kinerja waktu dan akurasi prediksi menjadi lebih singkat tetapi kinerja akurasi tetap baik. Karena hal tersebut maka kedua algoritma tersebut akan kita bandingkan dari tingkat keakurasiannya maupun kinerja waktu untuk meningkatkan kinerja prediksi. Langkah-langkah analisis pembiayaan dengan metode *Naive Bayes Classifier* adalah dengan menghitung jumlah class/label, menghitung jumlah kasus yang sama dengan class yang sama, mengalikan semua hasil variable, membandingkan hasil class dengan hasil terbesar akan dijadikan sebagai keputusan. Metode *Decision Tree (C4.5)* merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi atau pengelompokan dan bersifat prediktif. Langkah-langkah analisis pembiayaan dengan metode ini adalah dengan memilih atribut sebagai akar lalu membuat cabang untuk masing-masing nilai lanjutan dengan membagi kasus dalam cabang dan mengulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama. Hasil analisa kedua metode tersebut dibandingkan dengan melihat kemudahan penggunaan dan ketepatan hasil perhitungan masing-masing algoritma. Dari perhitungan yang telah di ujicoba hal tersebut menunjukkan bahwa metode *Decision Tree (C4.5)* memiliki tingkat akurasi lebih tinggi dan efisiensi waktu yang lebih cepat daripada metode *Naive Bayes Classifier*.

Kata Kunci: Pembiayaan, *Classifier*, *Decision Tree*, *C4.5*, *Nive Bayes*.

ABSTRACT

One of the classification algorithms are frequently used and got a lot of attention of researchers in predicting the financing problems in banking is Classifier Naive Bayes and Decision Tree (C4.5). But the disadvantages faced on both these algorithms is the duration and degree of prediction accuracy which is used to make predictions. This issue is also of concern to many researchers to fix it so that the performance time and the prediction accuracy is becoming shorter but performance remains good accuracy. Because of these conditions, the two algorithms are we going badingkan of rate Accuracy and performance prediction of time to improve performance. Step-by-step analysis of the financing method Clasifier Naive Bayes is to count the number of class / label, count the number of the same case with the same class, multiply all variable results, comparing the results with the results of the largest class will be used as a decision. Method of Decision Tree (C4.5) is one of the algorithms used to perform the classification or segmentation or clustering and predictive. Step-by-step analysis of the financing with this method is to select the attributes as root

and create a branch for each value continue to divide the case into a branch and repeat the process for each branch until all cases the branches have the same class. Results of analysis of both methods are compared to see the ease of use and accuracy of calculation of foreign respective algorithms. From calculations that have been in trials it shows that the method Decision Tree (C4.5) has a higher degree of accuracy and efficiency in a faster time than the method Naive Bayes classifier.

Keywords: *Financing, Classifier, Decision Tree, C4.5, Naive Bayes.*

I. PENDAHULUAN

Salah satu algoritma klasifikasi yang sering digunakan dan mendapat banyak perhatian para peneliti dalam memprediksi pembiayaan bermasalah pada perbankan adalah Naive Bayes Classifier dan Decision Tree (C4.5). Kesederhanaan pada algoritma Naive Bayes dan Decision Tree (C4.5) yang membuat kedua algoritma tersebut mempunyai daya tarik untuk diimplementasikan dalam berbagai aplikasi. Tetapi kelemahan yang dihadapi pada kedua algoritma tersebut adalah lamanya waktu dan tingkat akurasi prediksi yang digunakan untuk melakukan prediksi. Masalah ini juga menjadi perhatian banyak peneliti untuk memperbaikinya agar kinerja waktu dan akurasi prediksi menjadi lebih singkat tetapi kinerja akurasi tetap baik. Karena hal tersebut maka kedua algoritma tersebut akan kita bandingkan dari tingkat keakurasiannya maupun kinerja waktu untuk meningkatkan kinerja prediksi.

Selama ini proses penerimaan pengajuan pembiayaan dilakukan dengan teknik penilaian 5C (Character, Capacity, Capital, Collateral dan Condition) dengan menuliskan hasil data survey yang dimasukkan kedalam form operasional letter. Dengan cara ini seringkali menimbulkan beberapa permasalahan, antara lain adalah hasil yang kurang efektif karena data yang dimasukkan hanya berfokus pada salah satu calon nasabah saja, tidak melihat dari data pengalaman pembiayaan sebelumnya.

Dari permasalahan diatas dapat diambil alternatif dengan cara memanfaatkan teknik data *mining* dengan membandingkan 2 metode untuk identifikasi tumbuh kembang balita menggunakan metode *decision tree (c4.5)* dan *naive bayes*. Dengan harapan setelah diolah dengan data *mining* dapat membantu memprediksi kemungkinan pembiayaan macet.

1.1 Naive Bayes Classifier

Naïve Bayes Classifier merupakan sebuah metoda klasifikasi yang berakar pada teorema Bayes. Sebelum menjelaskan Naïve Bayes Classifier ini, akan dijelaskan terlebih dahulu Teorema Bayes yang menjadi dasar dari metoda tersebut. Pada teorema Bayes, bila terdapat dua kejadian yang terpisah (misalkan A dan B), maka teorema Bayes dirumuskan sebagai berikut:

$$P(A|B) = \frac{P(A)}{P(B)} P(B|A)$$

Teorema Bayes sering pula dikembangkan mengingat berlakunya hukum probabilitas total, menjadi seperti berikut:

$$P(A|B) = \frac{P(A)P(B|A)}{\sum_{i=1}^n P(A_i|B)}$$

dimana $A_1 \cup A_2 \cup \dots \cup A_n = S$

Untuk menjelaskan teorema Naïve Bayes, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema Bayes di atas disesuaikan sebagai berikut:

$$P(C|F_1, \dots, F_n) = \frac{P(C)P(F_1, \dots, F_n|C)}{P(F_1, \dots, F_n)}$$

Dimana variabel C merepresentasikan kelas, sementara variabel $F_1 \dots F_n$ merepresentasikan karakteristik-karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel dengan karakteristik tertentu dalam kelas C (posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut prior), dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas C (disebut juga likelihood), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga evidence). Karena itu, rumus (3) dapat pula ditulis secara sederhana sebagai berikut:

$$Posterior = \frac{prior \times likelihood}{evidence}$$

Nilai evidence selalu tetap untuk setiap kelas pada satu sampel.

Penjabaran lebih lanjut rumus Bayes tersebut dilakukan dengan menjabarkan $P(F_1, \dots, F_n|C)$ menggunakan aturan perkalian, menjadi sebagai berikut:

$$\begin{aligned} P(F_1, \dots, F_n|C) &= P(F_1|C)P(F_2, \dots, F_n|C, F_1) \\ &= P(F_1|C)P(F_2|C, F_1)P(F_3, \dots, F_n|C, F_1, F_2) \\ P(F_1, \dots, F_n|C) &= P(F_1|C)P(F_2|C, F_1) \dots P(F_n|C, F_1, F_2, \dots, F_{n-1}) \end{aligned}$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor-faktor syarat yang mempengaruhi nilai probabilitas,

Di sinilah digunakan asumsi independensi yang sangat tinggi (naïf), bahwa masing-masing petunjuk ($F_1, F_2 \dots F_n$) saling bebas (independen) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan sebagai berikut:

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i)$$

untuk $i \neq j$, sehingga

$$P(F_i|C, F_j) = P(F_i|C)$$

Dari persamaan di atas dapat disimpulkan bahwa asumsi independensi naïf tersebut membuat syarat peluang menjadi sederhana, sehingga perhitungan menjadi mungkin untuk

dilakukan. Selanjutnya, penjabaran $P(F_1, \dots, F_n | C)$ dapat disederhanakan menjadi seperti berikut:

$$P(F_1 \dots F_n | C) = P(F_1 | C) P(F_2 | C) \dots P(F_n | C)$$
$$P(F_1 \dots F_n | C) = \prod_{i=1}^n P(F_i | C)$$

Dengan kesamaan di atas, persamaan teorema Bayes dapat dituliskan sebagai berikut:

$$P(C | F_1 \dots F_n) = \frac{1}{P(F_1, F_2, \dots, F_n)} P(C) \prod_{i=1}^n P(F_i | C)$$
$$P(C | F_1 \dots F_n) = \frac{P(C)}{Z} \prod_{i=1}^n P(F_i | C)$$

Persamaan di atas merupakan model dari teorema Naïve Bayes yang selanjutnya akan digunakan dalam proses klasifikasi dokumen. Adapun Z merepresentasikan evidence yang nilainya konstan untuk semua kelas pada satu sampel.

1.2 Algoritma C4.5

Secara umum Algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

- Pilih atribut sebagai akar
- Buat cabang untuk masing-masing nilai
- Bagi kasus dalam cabang
- Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti yang tertera berikut:

Keterangan:

S : Himpunan kasus
 A : Atribut
 n : Jumlah partisi atribut A
 $|S_i|$: Jumlah kasus pada partisi ke i
 $|S|$: Jumlah kasus dalam S

Sebelum mendapatkan nilai Gain adalah dengan mencari nilai Entropi. Entropi digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan sebuah atribut. Rumus dasar dari Entropi adalah sebagai berikut:

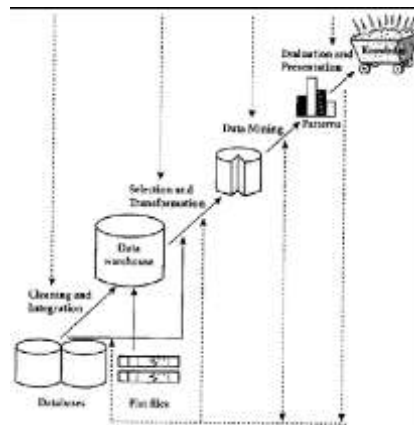
$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i$$

Keterangan:

S : Himpunan Kasus

n : Jumlah partisi S
 p_i : Proporsi dari S_i terhadap S

A. Tahap-tahap Data Mining



Gambar 1. Tahapan Data Mining

1. Pembersihan data (*Cleaning data*)

Untuk menghilangkan data yang tidak diperlukan, data yang diperoleh dari tahap pengambilan dataset akan disaring untuk menghasilkan data yang benar-benar dibutuhkan. umumnya data tersebut memiliki nilai yang tidak sempurna seperti data yang hilang. Selain itu, ada juga atribut-atribut data yang tidak sesuai dengan pemrosesan data mining yang akan digunakan. Data-data yang tidak relevan itu juga lebih baik dibuang karena keberadaannya bisa mengurangi mutu atau akurasi dari hasil data mining nantinya. Pembersihan data juga akan mempengaruhi performansi dari sistem data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2. Integrasi data

Data yang akan digunakan untuk data mining tidak hanya berasal dari satu database tetapi juga berasal dari beberapa database atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lain-lain. Pada tahap ini hal yang perlu dilakukan untuk lebih detail dan cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan keputusan pada akhirnya. Dalam integrasi data ini juga perlu dilakukan transformasi dan pembersihan data karena seringkali data dari dua database berbeda tidak sama cara penulisannya atau bahkan data yang ada di satu database ternyata tidak ada di database lainnya.

3. Seleksi Data

Data diseleksi untuk menentukan variabel apa saja yang akan diambil agar tidak terjadi kesamaan dan perulangan yang tidak diperlukan dalam pengolahan teknik data mining. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli

dalam kasus *market basket analysis*, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

4. Transformasi data

Pengubahan data menjadi format ekstensi yang sesuai untuk pengolahan dalam *data mining*. Beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa diproses dalam teknik data mining. Misalnya sebagian metode standar seperti analisis asosiasi dan klastering hanya bisa menerima input data kategorial. Karenanya data berupa angka numerik yang berlanjut perlu dibagimenjadi beberapa interval.

5. Proses *mining*

Untuk memproses teknik utama saat metode diterapkan agar menemukan pengetahuan berharga, data yang terkumpulkan sesuai prosedur harus di terapkan pada proses mining setelah data melalui tahap transformasi.

6. Evaluasi pola

Tahap ini yaitu mengidentifikasi pola-pola menarik kedalam *knowledge based* yang diidentifikasi. Dalam tahap ini hasil dari teknik *data mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah kajian yang ada sudah memenuhi target yang diinginkan. Jika ternyata hasil yang diperoleh tidak sesuai kajian ada beberapa alternatif dengan mencoba metode *data mining* lain agar lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

7. Presentasi pengetahuan

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahapterakhir dari proses data mining adalah bagaimana memformulasikan keputusanatau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil *datamining*.

B. *Confusion Matrix*

Pada data mining untuk mengukur atau ada bebeapa cara untuk mengukur kinerja dari model yang dihasilkan salah satunya menggunakan confusion matriks (akurasi) *Confusion matrix* adalah suatu metode yang digunakan untuk melakukan perhitungan akurasi pada konsep data mining. Presisi atau *confidence* adalah proporsi kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya. *Recall* atau *sensitivity* adalah proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar.

Tabel 2.1 Model *Confusion Matrix*

Aktual	Classified as	
	+	-
+	True positives (A)	False negatives (B)
-	False positives (C)	True negatives (D)

Perhitungan akurasi dengan tabel *confusion matrix* adalah sebagai berikut:

$$\text{Akurasi} = (A+D)/(A+B+C+D)$$

Presisi didefinisikan sebagai rasio item relevan yang dipilih terhadap semua item yang terpilih. Presisi dapat diartikan sebagai kecocokan antara permintaan informasi dengan jawaban terhadap permintaan tersebut. Rumus presisi adalah:

$$\text{Presisi} = A/(C+A)$$

Recall didefinisikan sebagai rasio dari item relevan yang dipilih terhadap total jumlah item relevan yang tersedia. *Recall* dihitung dengan rumus:

$$\text{Recall} = A/(A+D)$$

Presisi dan *Recall* dapat diberi nilai dalam bentuk angka dengan menggunakan perhitungan persentase (1-100%) atau dengan menggunakan bilangan antara 0-1. Sistem rekomendasi akan dianggap baik jika nilai presisi dan *recall*nya tinggi.

Kurva ROC menunjukkan akurasi dan membandingkan klasifikasi secara visual. ROC mengekspresikan *confusion matrix*. ROC adalah grafik dua dimensi dengan *false positive* sebagai garis horizontal dan *true positive* sebagai garis vertikal. AUC (*the area under curve*) dihitung untuk mengukur perbedaan performansi metode yang digunakan. ROC memiliki tingkat nilai diagnosa yaitu:

- a. Akurasi bernilai 0,90 – 1,00 = *excellent classification*
- b. Akurasi bernilai 0,80 – 0,90 = *good classification*
- c. Akurasi bernilai 0,70 – 0,80 = *fair classification*
- d. Akurasi bernilai 0,60 – 0,70 = *poor classification*
- e. Akurasi bernilai 0,50 – 0,60 = *failure*

2. METODE PENELITIAN

2.1. Metode Pengumpulan Data

Dalam penyusunan penelitian ini metode yang digunakan dalam proses pengumpulan data antara lain:

- a. Studi Pustaka
Studi pustaka dilakukan dengan cara menghimpun informasi yang relevan dengan topik atau masalah yang akan atau sedang diteliti. Informasi diperoleh dari buku-buku ilmiah, laporan penelitian, karangan-karangan ilmiah, tesis dan disertasi, peraturan-peraturan, ketetapan-ketetapan, ensiklopedia, dan sumber-sumber tertulis baik tercetak maupun elektronik.
- b. Dokumentasi
Pengambilan data dilakukan melalui dokumen tertulis maupun elektronik dari lembaga/institusi. Dokumen diperlukan untuk mendukung kelengkapan data yang lain.
- c. Wawancara
Wawancara merupakan proses interaksi atau komunikasi secara langsung antara pewawancara dengan responden. Wawancara dilakukan dengan cara mengajukan

pertanyaan sesuai dengan kebutuhan informasi peneliti. Informan dapat menyampaikan informasi yang komprehensif sebagaimana diharapkan peneliti.

A. Analisis Masalah

Semakin pesatnya penggunaan komputer dan alat-alat elektronik lainnya memacu pertumbuhan ketersediaan data, yang pada gilirannya memerlukan dukungan dari teknologi informasi untuk mengubah data tersebut menjadi suatu informasi, dan selanjutnya menjadi suatu pengetahuan yang bermanfaat. Seperti halnya analisis penerimaan pengajuan pembiayaan pada BMT yang perlu memanfaatkan teknologi informasi dalam hal ini dapat menggunakan komputer atau alat elektronik lainnya seperti smart phone, laptop atau komputer, agar dapat memudahkan karyawan BMT terutama pada bagian *account officer* guna ikut berpartisipasi dalam kemajuan BMT. Analisis penerimaan pengajuan pembiayaan sangat banyak pertimbangan yang diklasifikasikan menjadi beberapa bentuk pertimbangan yaitu jumlah pendapatan, jenis pembiayaan, status, jaminan, pekerjaan dan lainnya. Misalnya klasifikasi pada pekerjaan yang termasuk pedagang, pegawai, pengusaha, buruh adalah jenis pekerjaan.

Setiap pekerjaan memiliki tingkat kesulitan dan kendala tersendiri yang harus bisa dihadapi dengan cermat, begitupun dengan seorang *account officer* di BMT. Walaupun masih dalam lingkup mikro dan sederhana, seringkali ditemui berbagai kendala dalam proses analisis kelayakan hingga pengambilan keputusan pembiayaan khususnya di BMT, permasalahannya tersebut dapat menyebabkan angsuran kredit tidak lancar atau bahkan macet. Keberadaan BMT sangat dekat dengan masyarakat sehingga hubungan yang terjalin antara BMT dan nasabah lebih erat dibandingkan pada lembaga keuangan lainnya. Hal tersebut menjadi tantangan bagi *account officer* dan komite pembiayaan dalam menjaga profesionalisme kerja. Jangan sampai keputusan pembiayaan diambil berdasarkan hubungan pribadi maupun emosional.

B. Analisis Data

Untuk menjalankan sistem ini dibutuhkan masukan data yang dijadikan sebagai bahan perbandingan dengan data yang ada pada basis data. Adapun data yang digunakan pada penelitian ini diambil langsung dari BMT Al-fadhila cabang Sukarama Bandar Lampung. Adapun data yang diperlukan untuk penelitian ini yaitu data nasabah yang telah melakukan pembiayaan dimana sumber data didapat dari database komputer admin BMT dan tujuannya adalah untuk mencari kategori yang dapat mewakili isi dari kategori pembiayaan sehingga dapat dilakukan analisa keterhubungan antara seluruh kategori dengan hasil analisa.

Penelitian ini dilakukan untuk mengklasifikasi data nasabah dengan tujuan untuk melakukan pengkategorian berdasarkan jenis kemampuan nasabah, misalnya saja apakah pekerjaan pedagang, jumlah pembiayaan Rp.1.000.000, jenis pembiayaan murabahah, lama pembiayaan 115hari dan jaminan berupa BPKB Motor merupakan pembiayaan lancar atau macet.

3. HASIL DAN PEMBAHASAN

A. Data

Sebelum melakukan proses perhitungan algoritma, terlebih dahulu dilakukan pengumpulan data. Masalah yang akan diolah datanya dalam pembahasan ini adalah memprediksi data calon pemohon pembiayaan di KSPPS BMT Al-Fadhila. Data nasabah yang digunakan adalah 80 data nasabah yang sudah melakukan pembiayaan dari bulan desember 2015 sampai dengan bulan januari 2016. Atribut-atribut data yang digunakan adalah Pekerjaan, Jumlah Pembiayaan, Jenis Pembiayaan, Lama Pembiayaan dan Jaminan. Sumber data didapat dari database komputer admin BMT.

B. Perhitungan Data Mining

Berikut perhitungan manual naïve bayes dengan menggunakan data pembiayaan BMT pada tabel 4.1

Tabel 1. Data Tabel Pembiayaan

Pembiayaan					
Pekerjaan	Jumlah	Jenis	Lama	Jaminan	Kategori
Pedagang	1jt	Murabahah	30 hr	B. Nikah	Lancar
Pedagang	1jt	Murabahah	90 hr	Ijazah	Macet
Pedagang	1jt	Murabahah	115 hr	Ijazah	Macet
Pedagang	2jt	Murabahah	115 hr	B. Nikah	Lancar
Pedagang	1jt	Murabahah	115 hr	B. Nikah	Macet

C. Tingkat Akurasi Metode

Berdasarkan hasil percobaan yang sudah dilakukan untuk menghitung kinerja menggunakan metode Confusion Matrix hasilnya adalah sebagai berikut:

1. Naïve Bayes Classifier = 50%
2. Decision Tree (C4.5) = 56%

```

=== Confusion Matrix ===
      a  b  <-- classified as
33 16 |  a = Lancar
24  7 |  b = Macet
  
```

Gambar 2. Akurasi *Confusion Matrix* pada *Naïve Bayes Classifier*

=== Confusion Matrix ===

```
a  b  <-- classified as
42  7 | a = Lancar
28  3 | b = Macet
```

Gambar 3. Akurasi *Confusion Matrix* pada *Decision Tree (C4.5)*

Dari informasi tersebut, kemudian akan dilakukan proses perhitungan rata-rata persentase akurasi keberhasilan pada *confusion matrix* data training sebagai berikut:

$$\text{Akurasi} = \frac{\text{Banyaknya prediksi yang benar}}{\text{total banyaknya prediksi}} + \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

$$\text{Akurasi Naive Bayes} = \frac{33 + 7}{33 + 16 + 24 + 7} + \frac{40}{80} = 0.5$$

$$\text{Akurasi C4.5} = \frac{42 + 3}{42 + 7 + 28 + 3} + \frac{45}{80} = 0.56$$

Dari hasil perhitungan diatas dapat dilihat nilai akurasi data *testing* dengan menggunakan algoritma C4.5 adalah 56%.

4. KESIMPULAN DAN SARAN

4.1 KESIMPULAN

Berdasarkan pembahasan yang telah diuraikan, yaitu tentang analisis pembiayaan dengan metode *Naive Bayes Classifier* dan metode *Decision Tree (C4.5)* untuk memprediksi kelancaran pembiayaan di KSSPS BMT Al-Fadhila, maka dapat disimpulkan beberapa hal sebagai berikut:

1. Dalam melakukan prediksi tingkat ketepatan pembiayaan, dengan menggunakan pemodelan metode *Naive Bayes Classifier* didapatkan tingkat akurasi tertinggi sebesar 50%. Sedangkan pada pemodelan *Decision Tree (C4.5)* memperoleh rata-rata tingkat akurasi tertinggi sebesar 56%. Oleh karena itu dapat disimpulkan berdasarkan tingkat akurasi, bahwa pemodelan metode *Decision Tree (C4.5)* lebih baik dalam melakukan prediksi pembiayaan pada KSSPS / BMT Al-Fadhila.
2. Berdasarkan hasil perhitungan kedua metode tersebut tingkat nilai ROC adalah *failure*. Cara penyelesaian metode keduanya memiliki kelebihan dan kekurangan masing-masing. Pada Metode Algoritma Naive Bayes Classifier lebih mudah untuk digunakan karena hanya memiliki alur perhitungan yang tidak panjang sedangkan pada Metode Algoritma Decision Tree (C4.5) jika data diubah atau ditambah maka perhitungan akan memerlukan waktu yang lebih lama lagi.

4.2 SARAN

Adapun saran bagi penulis selanjutnya sebagai berikut:

1. Penelitian ini dapat dijadikan sebagai salah satu referensi bagi KSSPS BMT atau Koperasi Syari'ah untuk digunakan sebagai acuan dalam pengambilan keputusan untuk memprediksi kelancaran pembiayaan.

2. Dalam penelitian ini penulis hanya membahas perbandingan analisis antara metode Algoritma Naive Bayes Classifier dengan Decision Tree (C4.5). Oleh karena itu bagi peneliti selanjutnya dapat mengembangkan metode-metode analisis yang lain mengingat cakupan metode analisis yang cukup banyak serta dapat dikembangkan dengan mengaplikasikan pada bidang ilmu dan studi kasus yang berbeda.

REFERENSI

- [1] Hertanto Widodo dkk. 1999. PAS (Panduan Akuntansi Syariah) Panduan Praktis Operasional Baitul Maal Wat Tamwil (BMT). Bandung: Mizan.
- [2] Ibrahim Lubis. 1995. Ekonomi Islam Suatu Pengantar Jilid 2. Jakarta: Kalam Mulia.
- [3] Jazuli, dan Yadi Janwari. 2002. Lembaga lembaga Perekonomian Umat. Jakarta: PT. Raja Grafindo Persada.
- [4] Jefri. 2013. Implementasi Algoritma C4.5 Dalam Aplikasi Untuk Memprediksi Jumlah Mahasiswa Yang Mengulang Mata Kuliah Di STMIK AMIKOM. Yogyakarta : Yogyakarta.
- [5] Jiawei Han, Micheline Kamber. 2001. *Data Mining Concept and Techniques*.
- [6] Kasmir. 2001. Manajemen Perbankan. Jakarta : PT Raja Grafindo Persada.
- [7] Kusriani, 2008. Computing For Humanity Algoritma C4.5. Yogyakarta.
- [8] Kusriani, dan Emha Taufik Luthfi. 2009. Algoritma Data Mining. Yogyakarta: Andi Offset.
- [9] Muhammad. 2002. Lembaga-lembaga Keuangan Umat Kontemporer. Yogyakarta: UII Press.
- [10] M.Dawan Raharjo. 1999. Islam dan Transformasi Sosial-Ekonomi. Yogyakarta: Pustaka Pelajar.
- [11] Muhammad Ridwan. 2004. Manajemen Baitul Maal Wa Tamwil (BMT). Yogyakarta: UII Press.
- [12] Munawaroh Holisatul dkk. 2013 Perbandingan Algoritma ID3 Dan C5.0 Dalam Identifikasi Penjurusan Siswa SMA", Jurnal Sarjana Teknik Informatika , Vol. 1, No. 1, pp. 1-12, Juni.
- [13] Prasetyo, Eko. 2012. Data Mining : Konsep dan Aplikasi Menggunakan Matlab. Yogyakarta: CV. Andi Offset.
- [14] Syafi'i Antonio. 2001. Bank Syari'ah Dari Teori Ke Praktek. karta: Penerbit Gema Insani.